



Bijlage 1 - Associaties tussen woonafstand tot veehouderijen en gezondheidsproblemen: aanvullende details

Gert Jan Boender, Dominika Kalkowska, Thomas Hagenaars
(Wageningen Bioveterinary Research)

Lidwien Smit, Dick Heederik (IRAS-UU)

Christos Baliatsas, Joris IJzermans (NIVEL)

Bijlage bij het RIVM Rapport 2017-0062: 'Veehouderij en Gezondheid Omwonenden (aanvullende studies) - Analyse van gezondheidseffecten, risicofactoren en uitstoot van bio-aerosolen'

Bijlage 1

Associaties tussen woonafstand tot veehouderijen en gezondheidsproblemen: aanvullende details

In deze bijlage worden zowel de methode als de resultaten van hoofdstuk 2 uit het rapport in meer technisch detail beschreven.

Vraagstellingen

Doel van het onderzoek was om te toetsen of er associaties zijn tussen het wonen in de nabijheid van veehouderijen (met verschillende typen dieren) en het voorkomen van een risico op de volgende aandoeningen:

- longontsteking;
- bovenste-luchtweginfecties;
- inflammatoire darmaandoeningen;
- (als controle-aandoening) lage rugpijn.

Als er een associatie bleek te bestaan, werd de grootte en reikwijdte van de risicoverhoging onderzocht, en is het zogenaamde populatie attributief risico (PAF) bepaald. Het PAF drukt uit hoeveel procent van de gevallen voorkomen zouden worden als niemand in het studiegebied binnen de reikwijdte van de betreffende veehouderijen zou wonen.

Methode: Waarom kernel-analyses?

Regressieanalyse is de eerste stap om ruimtelijke associaties te onderzoeken en geeft inzicht in de mogelijke associaties tussen gezondheidsproblemen en grootheden die een maat vormen voor blootstelling aan veehouderij. In het VGO onderzoek (Maassen et al, 2016) worden grootheden als 'aantal varkens binnen een bepaalde afstand tot de woning', of 'afstand tot het dichtstbijzijnde pluimveebedrijf' gebruikt als maat voor verschillende typen blootstelling. Het regressiemodel beschrijft de mogelijke risicobronnen dus niet individueel maar samengevat. Als de regressieanalyse laat zien dat de afstand tot een dichtstbijzijnde veehouderij met een bepaald type dier een risicofactor is voor de gezondheid van omwonenden, zijn dergelijke veehouderijen mogelijk een bron van deze gezondheidslast. Door aanvullend een kernel-analyse uit te voeren kan extra informatie worden verkregen. Kernel-modellen beschrijven *per individuele (mogelijke) risico-bron* de kans op optreden van een gezondheidsprobleem bij omwonenden in afhankelijkheid van de afstand van de risicobron tot de woning. Een kernel-model voor gezondheidslast rondom veehouderijbedrijven geeft daarom – in aanvulling op regressiemodellen – inzicht in de bijdrage van ieder individueel veehouderijbedrijf aan een eventuele risicoverhoging voor een gezondheidsprobleem bij omwonenden. Kernel-modellen worden 'gefit' aan de gegevens door voor elke blootgestelde alle bijdragen van alle mogelijke bronnen (elk op basis van hun individuele afstand tot de blootgestelde) in de berekening mee te nemen. Dit vereist meer computerkracht en geheugenruimte dan noodzakelijk bij regressieanalyses.

Met het complete inzicht van regressie- en kernel-analyses kunnen vanuit gevonden associaties conclusies worden getrokken over aannemelijkheid van causaliteit. Omdat kernel-modellen de bijdrage aan gezondheidslast *per bedrijf* beschrijven, vormen ze tevens een basis om cumulatieve effecten van meerdere bedrijven te berekenen voor gegeven lokale situaties.

Methode: technische beschrijving

Gebruikte gegevens

De in dit onderzoek uitgevoerde kernel-analyses maken gebruik van de volgende gegevens over de jaren 2009 t/m 2013:

1. Gegevens over de incidentie (of in sommige gevallen prevalentie) van gezondheidsproblemen bij patiënten van de deelnemende huisartsenpraktijken;
2. Afstanden tussen de woonadressen van die patiënten en alle veehouderijen in het studiegebied. Voor berekening van deze afstanden werden locatiecoördinaten van alle veehouderijen in het studiegebied verkregen uit de openbare provinciale databases van vergunningen voor het houden van vee (BVB, bestand veehouderijbedrijven, mei 2012), en coördinaten van woonlocaties van huisartsenpatiënten door geo-coderen van woonadressen.

Om de privacy van patiënten te waarborgen is gewerkt met bestanden waarin patiënten geanonimiseerd zijn opgenomen. De gegevens (2) betroffen afstanden in plaats van woonlocaties zodat (1) en (2) gekoppeld konden worden zonder dat gezondheidsinformatie van de geanonimiseerde patiënten gekoppeld zou kunnen worden aan hun woonlocatie. Daarbij werden de voor (2) berekende afstanden zodanig afgerond dat de woonlocatie van de patiënt daaruit niet weer terug berekend zou kunnen worden.

Voor de analyse van huisartsengegevens is gebruik gemaakt van gegevens uit de elektronische patiëntendossiers (EPD's) van de deelnemende huisartsenpraktijken. Deze gegevens zijn eerder beschreven in het VGO rapport (paragraaf 2.5). Huisartsenpraktijken werden geselecteerd op vooraf bepaalde registratiekwaliteitscriteria (Heederik en IJzermans, 2011; Smit et al., 2012, van Dijk et al. 2017). Medische gegevens van 110.728 patiënten uit 27 huisartsenpraktijken in het VGO-studiegebied die aan de criteria voldeden konden gebruikt worden voor analyse van de "VGO" periode 2010-2013. Ook voor 2009 konden gegevens van patiënten uit huisartspraktijken in het VGO-studiegebied die aan dezelfde criteria voldeden, worden gebruikt. Dit ging om 22 huisartspraktijken waarvan er 17 ook voorkwamen in de gebruikte gegevens voor de periode 2010-2013. Het totaal aantal verschillende betrokken huisartspraktijken over de periode 2009-2013 kwam daarmee op 32. Er is geen restrictie op leeftijd toegepast, d.w.z. de analyses zijn uitgevoerd voor patiënten van alle leeftijden. Dit in tegenstelling tot de eerdere analyses van longontsteking voor 2009 zoals gepresenteerd in het hoofdrapport en in het artikel van Smit et al. (2017), die op basis van patiënten van 0-70 jaar waren. Omdat longontsteking bij ouderen veel vaker voorkomt dan bij jongeren, leidt dit er toe dat het cijfer voor het (gemiddeld) aantal gevallen van longontsteking per 1000 patiënten in 2009 zoals gegeven in Tabel 2.1 duidelijk hoger is dan het cijfer in het artikel van Smit et al. (2017).

Alleen de gegevens van huisartsenpraktijken die in een bepaald jaar aan een aantal strikte eisen voldeden werden in de analyses opgenomen. Algemene eisen per jaar voor inclusie van een huisartsenpraktijk voor de incidentie- en prevalentiecijfers zijn als volgt:

- Bij minimaal 50% van de morbiditeitsrecords of episoderecords moet een ICPC-code zijn geregistreerd.
- Het percentage records per kwartaal van zowel morbiditeits- als prescriptierecords moet minimaal 10% zijn.

Extra eis voor incidentiecijfers:

- In 75% of meer van de morbiditeitsrecords dient een 'aard van het contact' te zijn geregistreerd of in minimaal 75% van de episodes moet een startdatum bekend zijn.

Voor de analyse van verergeringen van COPD werden de volgende gezondheidsgegevens gebruikt. In 21 deelnemende praktijken werden patiënten via de huisarts uitgenodigd een korte enquête in te vullen over mogelijke luchtwegproblemen (Borlée et al., 2015a). Om in aanmerking te komen moesten de patiënten: a) in het oostelijk deel van Noord-Brabant of het noorden van Limburg wonen; b) wonen in een gemeente met minder dan 30.000 inwoners; en c) tussen 18-70 jaar oud zijn. Uit de in aanmerking komende patiënten, werd per huisadres één persoon willekeurig gekozen. Van de 27.869 uitgenodigde personen die een vragenlijst hebben ontvangen hebben er 14.882 de vragenlijst ingevuld teruggestuurd (respons van 53,4%). Patiënten met ontbrekende gegevens of zonder een elektronisch patiëntendossier (EPD), mensen die wonen of werken op een boerderij met dieren en mensen die minder dan één jaar op hun huidige huisadres wonen werden uitgesloten, waarna er 12.117 vragenlijsten in aanmerking kwamen voor analyse (Borlée et al., 2015a; Borlée et al., 2015b).

De coördinaten van alle veehouderijbedrijven in het studiegebied zijn verkregen uit de provinciale databank voor verplichte milieuvergunningen voor het houden van vee in 2009 en 2012 (Bestand Veehouderij Bedrijven, BVB) van de provincies Noord-Brabant en Limburg. Ook de diersoort (hoofdcategorie) en het aantal dieren per bedrijf waren in deze database beschikbaar. Voor de kernel-analyses is het type veehouderij gebaseerd op de hoofdcategorie van de diersoort, waarbij voor geitenhouderijen de aanvullende eis werd opgelegd van minimaal 50 dieren.

Statistische analyse

De kernel beschrijft een kansfunctie, voor individuele patiënten op het ervaren van een gezondheidsprobleem in een gegeven periode, onafhankelijk gegenereerd door elk veehouderijbedrijf van een bepaald type [Smit et al 2017]. Deze kansfunctie wordt verondersteld afhankelijk te zijn van de afstand tussen de veehouderijbedrijven en de woonlocatie: $p = p(r_{ij})$, waar r_{ij} de hemelsbrede afstand is tussen bedrijf i en de woonlocatie van individu j . We gebruiken de herschrijving $p(r) = 1 - \exp(-\lambda(r))$ waarin $\lambda(r)$ de zogenaamde hazardfunctie is. In de analyses hier uitgevoerd wordt de hazardfunctie als een stapfunctie geparameteriseerd:

$$\begin{aligned}\lambda(r) &= \lambda_0 \text{ als } r \leq d; \\ \lambda(r) &= 0 \text{ als } r > d.\end{aligned}$$

Hierbij is λ_0 een constante die aan de gegevens wordt gefit, en d een vooraf gekozen waarde voor de reikwijdte.

We introduceren een afstands-onafhankelijke achtergrondkans P_b , zodat die component van het gezondheidsprobleem dat niet geassocieerd is met nabijheid tot veehouderijbedrijven in de berekening kan worden betrokken. Equivalent met deze achtergrondkans definiëren we een afstands-onafhankelijke achtergrondhazard

$\lambda_b = \ln(1 - P_b)$. Deze achtergrondhazard λ_b en de kernelparameter λ_0 (en eventuele verdere kernelparameters) worden met behulp van Maximum-Likelihood (ML) geschat. We berekenen 95% betrouwbaarheidsgrenzen voor de parameters met behulp van de likelihood-ratio toets. De likelihoodfunctie wordt gegeven door

$$L = \prod_{j \in \Lambda^{\text{NP}}} p_{\text{esc},j} \prod_{m \in \Lambda^{\text{P}}} (1 - p_{\text{esc},m})$$

Hier staan $p_{\text{esc},j}$ (resp. $p_{\text{esc},m}$) voor de (totale) kans dat een individu j (resp. m) ontsnapt aan een gezondheidsprobleem, Λ^{NP} is de groep van alle patiënten in de studiepopulatie zonder het gezondheidsprobleem, en Λ^{P} is de groep van alle patiënten in de studiepopulatie met het gezondheidsprobleem. Uit de veronderstelde onafhankelijkheid van de door elke individuele veehouderij gegenereerde kans op het gezondheidsprobleem volgt dat de totale kans $p_{\text{esc},j}$ dat een individu j ontsnapt aan het gezondheidsprobleem gelijk is aan $p_{\text{esc},j} = (1 - P_b) \prod_i p_{\text{esc},ji}$, waar de index i loopt over alle veehouderijen van het betreffende type in het studiegebied, $1 - P_b$ is de kans dat een individu ontsnapt aan het gezondheidsprobleem voor zover veroorzaakt door de achtergrondhazard, en $p_{\text{esc},ji} = 1 - p(r_{ij})$ is de kans dat een individu j ontsnapt aan het gezondheidsprobleem voor zover veroorzaakt door veehouderij i . Gebruik makend van $p(r) = 1 - \exp(-\lambda(r))$, kunnen we schrijven: $p_{\text{esc},ji} = \exp(-\lambda(r_{ij}))$, en dus kan de likelihood als volgt geschreven worden in termen van $\lambda(r)$:

$$L = \prod_{j \in \Lambda^{\text{NP}}} \exp\left(-\lambda_b - \sum_i \lambda(r_{ij})\right) \prod_{m \in \Lambda^{\text{P}}} \left(1 - \exp\left(-\lambda_b - \sum_i \lambda(r_{im})\right)\right).$$

Resultaten en discussie

Analyse van longontsteking in de huisartspraktijk met stapfuncties rond alle bedrijfstypen

In de multivariate analyses per jaar is voor elk bedrijfstype de 'best-fit' reikwijdte genomen uit de individuele analyse, binnen de range waarvoor een consistent signaal gevonden werd. Voor de volledigheid zijn ook multivariate analyses gedaan met andere combinaties van reikwijdtes. De resultaten hiervan duiden er op dat de andere combinaties geen betere fit opleveren. In Tabel 1 staan voor de volledigheid vermeld de ranges van reikwijdtes waarvoor in de individuele analyses rondom pluimvee- en geitenhouderijen een significant verhoogd risico op longontsteking werd gevonden.

Tabel 1. Ranges van reikwijdtes (in km) waarvoor in de individuele analyses rondom pluimvee- en geitenhouderijen een significant verhoogd risico op longontsteking werd gevonden.

Jaar	Geiten	Pluimvee
2009	0,5–5	0,5–1,5
2010	1,5–2	0,5–1,5
2011	1–3,5	1–1
2012	1–5	0,5–1,5
2013	1–5	1–1,5

In de individuele analyse worden voor rundveebedrijven in vier van de vijf jaren associaties met longontsteking in omwonenden gevonden, maar deze verdwijnen grotendeels in de multivariate analyse. Dit illustreert dat wanneer de effecten van verschillende bedrijfstypen voor elkaar worden gecorrigeerd, een deel van die effecten kan verdwijnen.

Analyse van longontsteking per huisartsenpraktijk

Univariate analyses per huisartsenpraktijk werden uitgevoerd voor longontsteking in relatie tot pluimvee en geiten, waarbij de gehele periode 2009-2013 bij elkaar genomen is. Deze analyses dienen om na te gaan of de uitkomsten van de overall analyses niet toe te schrijven zijn aan een lokaal effect in bijvoorbeeld één huisartsenpraktijk. Als voor meerdere huisartsenpraktijken apart een associatie wordt gevonden geeft dat bovendien verder vertrouwen dat de associaties gevonden in de overall analyses niet berusten op een artefact dat misschien zou kunnen ontstaan door verschillen in geregistreerde incidentie tussen huisartsen(praktijken). Voor longontsteking rond pluimveebedrijven vinden we in 9 van de 32 praktijken een significante verhoging van longontsteking; voor 5 van de 9 praktijken treedt een significant effect op bij meer dan één reikwijdte. De 9 praktijken liggen verspreid over het VGO gebied, hoewel niet gelijkmatig verspreid. Voor longontsteking rond geitenbedrijven vinden we in 12 van de 32 praktijken een significante verhoging; voor 3 van de 12 praktijken treedt een significant effect op bij meer dan één gekozen reikwijdte. Aangezien analyses per praktijk minder onderscheidend vermogen hebben, is het niet verrassend dat de associatie niet voor alle praktijken afzonderlijk gevonden wordt. Het feit dat deze voor 9, respectievelijk 12, van

de 32 praktijken wel gevonden wordt, laat zien dat de associatie niet op een klein deelgebied berust.

Analyse van longontsteking rondom geitenbedrijven met onderscheid tussen bedrijven naar Q-koortsstatus

Om na te gaan of de gevonden associatie tussen risicoverhoging voor longontsteking en nabijheid tot geitenbedrijven niet uitsluitend toe te schrijven is aan bedrijven die eerder besmet zijn geweest met Q-koorts, is een kernel-analyse gedaan met onderscheid tussen geitenbedrijven naar Q-koortsstatus. Onderscheiden zijn bedrijven die in de periode 2006-2013 nooit de status van besmet bedrijf (met Q-koorts) kregen, en bedrijven die in die periode op enig moment de status van besmet bedrijf hadden. De laatste zijn op hun beurt opgedeeld in besmette bedrijven met abortusgolf, en besmette bedrijven zonder (gemelde) abortusgolf (alleen tankmelk positief). Ook de pluimveebedrijven zijn opgenomen in de analyse. De resultaten zijn weergegeven in Tabel 2. In vier van de vijf geanalyseerde jaren wordt een associatie gevonden tussen longontsteking en nabijheid tot geitenbedrijven die in de periode 2006-2013 nooit de status van besmet bedrijf hadden. Dit laat zien dat de gevonden associatie tussen risicoverhoging voor longontsteking en nabijheid tot geitenbedrijven niet uitsluitend toe te schrijven is aan bedrijven die eerder besmet zijn geweest met Q-koorts.

Tabel 2. Resultaten van multivariate kernel-analyses voor longontsteking rondom geitenbedrijven met onderscheid tussen bedrijven naar Q-koortsstatus in de periode 2006-2013. Q-abortus: besmette bedrijven met abortusgolf; Q-tankmelk: besmette bedrijven zonder (gemelde) abortusgolf; Q-negatief: nooit aangemerkt als besmet bedrijf. +: associatie tussen risicoverhoging en nabijheid tot type bedrijf; -: geen associatie.

	Q-abortus	Q-tankmelk	Q-negatief	Pluimvee
2009	+	+	+	+
2010	-	-	+	+
2011	+	+	+	+
2012	+	+	+	+
2013	+	+	-	+

Kernel-analyse voor longontsteking met algemenere kansfuncties

In de eerdere analyse van IVG longontsteking data uit 2009 (zie VGO hoofdrapport en Smit et al 2017) is gebruik gemaakt van de algemene kansfunctie met drie te fitten parameters. De best-fit kansfunctie neemt de gedaante van een stapfunctie aan; bevestigend dat er een significant extra risico op longontsteking is rondom pluimveebedrijven, en suggererend dat die vrij scherp afneemt boven een afstand van ca. 1.15km. Wanneer eenzelfde aanpak wordt gevolgd bij de analyse van (een deel van) de VGO huisartsengegevens voor longontsteking (met beperking tot pluimvee- en geitenbedrijven als mogelijke bronnen) wordt ook een stapfunctie als beste fit gevonden. Om te zien of de stapfunctie optreedt (in plaats van een meer vloeiend verlopende functie) doordat de parametrisatie niet genoeg vrijheid geeft, is een set alternatieven bestudeerd. Echter alle alternatieven wijzen er steeds op dat de stapfunctie de beste fit

geeft. Vervolgens zijn de gegevens voor longontsteking in relatie tot pluimveebedrijven ook geanalyseerd met een stapfunctie-parametrisatie waarbij de reikwijdte niet vooraf werd gekozen maar werd mee-gefit. De best-fit reikwijdte waarbinnen het risico verhoogd is verschilt enigszins tussen de jaren en beweegt zich tussen waarden van iets minder dan 1 tot ongeveer 2 km. Voor 2012 en 2013 is de onzekerheidsmarge op de reikwijdte relatief groot in verschillende groepen (bij opdeling naar geslacht en leeftijd) en lijken verschillende reikwijdtes op te treden. Of deze beide aspecten van stapfunctie en variatie in reikwijdte verklaard kunnen worden uit blootstellingspatronen is op dit moment onduidelijk. Verder onderzoek zou hier antwoord op kunnen geven. Bij het zoeken naar onderliggende verklaringen voor de gevonden associaties is het ook van belang om een eventuele bedrijfsgrootte-afhankelijkheid van de gevonden associatie in kaart te brengen. Voor longontsteking rondom geitenbedrijven was het mogelijk een bedrijfsgrootte-afhankelijk model te fitten, waarbij de reikwijdte afhankelijk is gekozen van de bedrijfsgrootte. Dit geeft een betere fit dan met één vaste reikwijdte voor alle bedrijfsgroottes. Deze rekenintensieve analyse is echter nog maar voor één enkel jaar uit de dataset uitgevoerd. Verder onderzoek is noodzakelijk om deze eerste bevinding uit te breiden met resultaten voor meerdere jaren.

Referenties

Boender GJ, Hagenaars TJ, Bouma A, Nodelijk G, Elbers ARW, De Jong MCM, Van Boven M (2007). Risk maps for the spread of highly pathogenic avian influenza in poultry. *PLoS Comput Biol* 3: e71.

Borlée F, Yzermans CJ, van Dijk CE, Heederik D, Smit LAM (2015). Increased respiratory symptoms in COPD patients living in the vicinity of livestock farms. *European Respiratory Journal*, 46: 1605-1614.

Borlée F, Yzermans CJ, Krop E, Aalders B, Rooijackers J, Zock JP, van Dijk CE, Maassen K, Schellevis F, Heederik D, Smit LAM (2017). Spirometry, questionnaire and Electronic Medical Record based COPD in a population survey: comparing prevalence, level of agreement and associations with potential risk factors. *PLoS ONE*; in press.

Van Dijk CE, Zock J-P, Baliatsas C, Smit LAM, Borlée F, Spreeuwenberg P, Heederik D, Yzermans CJ (2017). Health conditions in rural areas with high livestock density: Analysis of seven consecutive years. *Environmental Pollution* 222, 374e382.

Heederik D, Yzermans J (2011). Mogelijke effecten van intensieve veehouderij op de gezondheid van omwonenden. Utrecht.

Maassen K, Smit L, Wouters I, Van Duijkeren E, Janse I, Hagenaars T, IJzermans J, Van der Hoek W, Heederik D (2016). Veehouderij en gezondheid omwonenden, Bilthoven.

Smit LAM, van der Sman-de Beer F, Opstal-van Winden AW, Hooiveld M, Beekhuizen J, Wouters IM, Yzermans J, Heederik D (2012). Q Fever and pneumonia in an area with a high livestock density: a large population-based study. *PLoS ONE* 7: e38843.

Smit LAM, Boender GJ, De Steenhuijsen Piters WAA, Hagenaars TJ, Huijskens EGW, Rossen JWA, Koopmans M, Nodelijk G, Sanders EAM, Yzermans J, Bogaert D, Heederik D (2017). Increased pneumonia risk near poultry farms: does the respiratory microbiota play a role? *Pneumonia* 9:3.